

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2001, 11:10:54 ; Search time 47.29 Seconds
(without alignments)
37.177 Million cell updates/sec

Title: US-09-288-719-3
Perfect score: 81
Sequence: 1 GGGSGGGRASGGGS 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database :
1: SP:REMBL.15:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.rodent:*
13: sp.virus:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	86.4	298	11 Q9QYF0	Q9QYF0 mus musculu
2	70	86.4	738	5 O02402	O02402 pinctada fu
3	69	85.2	150	10 O80490	O80490 arabidopsis
4	67	82.7	586	5 O9NAJ4	O9NAJ4 caenorhabdi
5	67	82.7	877	4 O9UP51	O9UP51 homo sapien
6	67	82.7	989	4 O9P271	O9P271 homo sapien
7	67	82.7	1433	11 O07563	O07563 mus musculu
8	67	82.7	1487	4 O9ULM3	O9ULM3 homo sapien
9	66	81.5	290	4 O9UTJ3	O9UTJ3 homo sapien
10	66	81.5	291	4 O14621	O14621 homo sapien
11	66	81.5	306	4 O9UKM9	O9UKM9 homo sapien
12	66	81.5	509	5 O96754	O96754 branchiosto
13	66	81.5	584	5 O45030	O45030 drosophila
14	66	81.5	592	2 O9P600	O9P600 xylella fas
15	66	81.5	944	5 O9VKS9	O9VKS9 drosophila
16	65	80.2	96	5 O9VKS2	O9VKS2 drosophila
17	65	80.2	133	10 O9LH30	O9LH30 oryza sativ
18	65	80.2	136	5 O18444	O18444 caenorhabdi
19	65	80.2	175	10 O9LSN6	O9LSN6 arabidopsis

20	65	80.2	221	10 O65514	O65514 arabidopsis
21	65	80.2	251	10 O9SNW4	O9SNW4 oryza sativ
22	65	80.2	301	5 O15647	O15647 plasmodium
23	65	80.2	316	5 O19111	O19111 caenorhabdi
24	65	80.2	318	10 O38777	O38777 allium sativ
25	65	80.2	344	13 O42403	O42403 gallus gall
26	65	80.2	384	10 O9LQW3	O9LQW3 arabidopsis
27	65	80.2	493	4 O14720	O14720 homo sapien
28	64	79.0	160	10 O9SMA8	O9SMA8 glycine max
29	64	79.0	376	5 O76880	O76880 drosophila
30	64	79.0	376	5 O9W4X2	O9W4X2 drosophila
31	63	77.8	155	5 O9VZK6	O9VZK6 drosophila
32	63	77.8	237	5 O9WZL8	O9WZL8 drosophila
33	63	77.8	239	5 O24345	O24345 drosophila
34	63	77.8	271	10 O08529	O08529 nicotiana t
35	63	77.8	312	5 O18097	O18097 caenorhabdi
36	63	77.8	344	11 O9JHT5	O9JHT5 mus musculu
37	63	77.8	706	4 O9UI36	O9UI36 homo sapien
38	63	77.8	876	5 O9VX88	O9VX88 drosophila
39	63	77.8	1180	4 O92625	O92625 homo sapien
40	63	77.8	1417	5 O9W3A4	O9W3A4 drosophila
41	63	77.8	1431	11 O9JMH4	O9JMH4 mesocricetu
42	63	77.8	3726	11 O61329	O61329 mus musculu
43	62	76.5	80	10 O9SUF7	O9SUF7 arabidopsis
44	62	76.5	118	5 O9VYS6	O9VYS6 drosophila
45	62	76.5	130	5 O9VYD8	O9VYD8 drosophila

ALIGNMENTS

RESULT 1
ID Q9QYF0 PRELIMINARY; PRT; 298 AA.
AC Q9QYF0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 13, Last annotation update)
DE CN 8 SCFV.
GN CN 8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=SPLEEN;
RA Shinozawa N., Demura T., Fukuda H.;
RN Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=SPLEEN;
RA Shinozawa N., Demura T., Fukuda H.;
RN Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[2]
RC STRAIN=BALB/C; TISSUE=SPLEEN;
RA Shinozawa N., Demura T., Fukuda H.;
RN Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[2]
RT "Isolation of a novel type of vascular cell wall-specific monoclonal
RT antibody recognizing a cell polarity using a phage display subtraction
RT method."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB036341; BAF86533.1; -
DR HSSP: P01607; IRET.
DR INTERPRO: IPR003006; -
DR PFM: PFM0047; 19; 2.
SQ SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

Query Match 86.4%; Score 70; DB 11; Length 298;
Best Local Similarity 86.7%; Pred. No. 0.15;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGSGGGRASGGGS 15
DB 158 GGGSGGGRASGGGS 172

```

RESULT 2
002402 PRELIMINARY; PRT; 738 AA.
AC 002402;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE INSOLUBLE PROTEIN.
OS Pinctada fucata.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pterioidea;
OC Pterioidea; Pteridae; Pinctada.
OX NCBI_TaxID=50426;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97320490; PubMed=9177341;
RA Sudo S., Fujikawa T., Nagakura T., Ohkubo T., Sakaguchi K., Tanaka M.,
RA Nakashima K., Takahashi T.;
RT "Structures of mollusc shell framework proteins."
RL Nature 387:563-564(1997).
DR EMBL: D86074; BAA20466.1;
SQ SEQUENCE 738 AA; 61723 MW; FDF984139BF3BA59 CRC64;

Query Match
Best Local Similarity 86.4%; Score 70; DB 5; Length 738;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGSGGGRASGGGS 15
Db 666 GGGSGGGRASGGGS 680

RESULT 3
080490 PRELIMINARY; PRT; 150 AA.
AC 080490;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE T12M4.13 PROTEIN.
GN T12M4.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLOMBIA;
RA Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Oj1, O, Kwan, A.,
RA Liu S., Li J., Arujo R., Au M., Brendel V., Buehler E., Conway A.B.,
RA Conway A.R., Dewar K., Feng J., Kim C., Kurtz D., Li Y., Palm C.J.,
RA Shim P., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.;
RT "Arabidopsis thaliana chromosome 1 BAC T12M4 sequence, complete
sequence."
RN Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLOMBIA;
RA Theologis A.;
RN Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RL [3]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLOMBIA;
RA Theologis A.;
RN Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RL [4]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLOMBIA;
RA Theologis A.;
RN Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC003114; AAC24095.1;

```

```

SQ SEQUENCE 150 AA; 15060 MW; C601BEAA65C170AD CRC64;

Query Match
Best Local Similarity 85.2%; Score 69; DB 10; Length 150;
Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGSGGGRASGGGS 15
Db 67 GGGSGGGRASGGGS 81

RESULT 4
09NAJ4 PRELIMINARY; PRT; 586 AA.
AC 09NAJ4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Y38E10A.17 PROTEIN.
GN Y38E10A.17.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wallis J.M.;
RN Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: AL110484; CAB54608.1;
SQ SEQUENCE 586 AA; 59942 MW; CC1B598208A3932A CRC64;

Query Match
Best Local Similarity 82.7%; Score 67; DB 5; Length 586;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGSGGGRASGGGS 15
Db 461 GGGSGGGRASGGGS 475

RESULT 5
09UF51 PRELIMINARY; PRT; 877 AA.
AC 09UF51;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE HYPOTHETICAL 94.9 KDA PROTEIN (FRAGMENT).
GN DKFP434P0917.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=TESTIS;
RA Bioecker H., Boecker M., Brandt P., Mewes H.W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL133612; CAB63741.1;
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 877 AA; 94937 MW; 8D024C40BFD5F82 CRC64;

```

Query Match 82.7%; Score 67; DB 4; Length 877;
Best Local Similarity 80.0%; Pred. No. 0.99;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGSGGGRASGGGS 15
DB 143 GGGGGGGSGGGGS 157

RESULT 6

09P271 PRELIMINARY; PRT; 989 AA.
AC 09P271;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE KIAA1457 PROTEIN (FRAGMENT).
GN KIAA1457.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human
genes.XVII The complete sequences of 100 new cDNA clones from brain
which code for large proteins in vitro."
RL DNA Res. 7:143-150(2000).
DR EMBL; AB040890; BAA95981.1; -.
FT NON_TER
SQ SEQUENCE 989 AA; 107305 MW; 12C5EDF6442569C8 CRC64;

Query Match 82.7%; Score 67; DB 4; Length 989;
Best Local Similarity 80.0%; Pred. No. 1.1;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGSGGGRASGGGS 15
DB 261 GGGGGGGSGGGGS 275

RESULT 7

007563 PRELIMINARY; PRT; 1433 AA.
AC 007563;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE PROCOLLAGEN, TYPE XVII, ALPHA 1 (BULLOUS PEMPHIGOID AUTOANTIGEN
BP180) (COLLAGEN XVII).
GN COL17A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/C;
RC MEDLINE=93232041; PubMed=8473327;
RA Li K., Tamai K., Tan E.M., Uitto J.;
RT "Cloning of type XVII collagen. Complementary and genomic DNA
sequences of mouse 180-kilodalton bullous pemphigoid antigen (BPAG2)
predict an interrupted collagenous domain, a transmembrane segment,
and unusual features in the 5'-end of the gene and the 3'-untranslated
region of the mRNA."
RL J. Biol. Chem. 268:8825-8834(1993).
CC -1- FUNCTION: THE COLLAGEN DOMAINS OF BP180 ALLOW THE HEMIDESMOSE TO
CC FORM STABLE INTERACTIONS WITH THE CONSTITUENTS OF THE EXTRACELLULAR
CC MATRIX OF THE CUTANEOUS BASEMENT MEMBRANE ZONE.
CC -1- SUBCELLULAR LOCATION: A COMPONENT OF THE HEMIDESMOSE.
CC -1- DISEASE: UPON DISRUPTION OF BP180 BY AUTOANTIBODIES THE

CC HEMIDESMOSE IS DISRUPTED LEADING TO THE BLISTERING SKIN DISORDER
CC BULLOUS PEMPHIGOID (BP).
DR EMBL; L08407; AAA37443.1; -.
DR MGI; MGI:88450; Coll17a1.
DR INTERPRO: IPR000087; -.
DR PIRAM; PFI01391; Collagen; 5.
DR Antigen; Cell adhesion.
SQ SEQUENCE 1433 AA; 144087 MW; B8F808832A19922C CRC64;

Query Match 82.7%; Score 67; DB 11; Length 1433;
Best Local Similarity 80.0%; Pred. No. 1.6;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGSGGGRASGGGS 15
DB 443 GGGGGGGSGGGGA 457

RESULT 8

09ULM3 PRELIMINARY; PRT; 1487 AA.
AC 09ULM3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE KIAA1197 PROTEIN (FRAGMENT).
GN KIAA1197.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Nagase T., Ishikawa K., Kikuno R., Hirosewa M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro."
RL DNA Res. 6:337-345(1999).
DR EMBL; AB033023; BAA86511.1; -.
FT NON_TER
SQ SEQUENCE 1487 AA; 157164 MW; 9DC4FC22CCCF4414 CRC64;

Query Match 82.7%; Score 67; DB 4; Length 1487;
Best Local Similarity 80.0%; Pred. No. 1.7;
Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGSGGGRASGGGS 15
DB 873 GGGGGGGSGGGGS 887

RESULT 9

09UJ3 PRELIMINARY; PRT; 290 AA.
AC 09UJ3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE D64K7.3 (AUTOANTIGEN P542).
GN D64K7.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031668; CAB43742.1; -.
DR INTERPRO; IPR000504; -.

DR PFAM: PF00076; trm; 1.
DR PROSITE: PS00030; RNP_1; UNKNOWN_1.
SQ SEQUENCE 290 AA; 30364 MW; 228F7E277DB066C9 CRC64;

Query Match 81.5%; Score 66; DB 4; Length 290;
Best Local Similarity 85.7%; Pred. No. 0.44;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGSGGSRASGGG 14
DB 219 GGGSGGSGGGGGG 232

RESULT 10
ID 014621 PRELIMINARY; PRT; 291 AA.

AC 014621;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE AUOANRIGEN P542.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=95190029; PubMed=7533788;
RA Vaughan J.H., Valbracht J.R., Nguyen M.D., Handley H.H., Smith R.S.,
RA Patrick K., Rhodes G.H.;
RT "Epstein-Barr virus-induced autoimmune responses. I. Immunoglobulin M
RT autoantibodies to proteins mimicking and not mimicking Epstein-Barr
RT virus nuclear antigen-1.";
RL J. Clin. Invest. 95:1306-1315(1995).
RN [2]
RP SEQUENCE FROM N.A.

RX MEDLINE=98018738; PubMed=9376072;
RA Rhodes G.H., Valbracht J.R., Nguyen M.-D., Vaughan J.H.;
RT "The p342 gene encodes an autoantigen that cross-reacts with EBNA-1 of
RT the Epstein Barr virus and which may be a heterogeneous nuclear
RT ribonucleoprotein.";
RL J. Autoimmun. 10:447-454(1997).
RN [3]
RP SEQUENCE FROM N.A.

RA Vaughan J.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.

RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; U38696; AAC28898.1; -.
DR INTERPRO: IPR000504; -.
DR PFAM: PF00076; trm; 1.
DR PROSITE: PS00030; RNP_1; UNKNOWN_1.
SQ SEQUENCE 291 AA; 30438 MW; F8DB4BE6DA08347F CRC64;

Query Match 81.5%; Score 66; DB 4; Length 291;
Best Local Similarity 85.7%; Pred. No. 0.44;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGSGGSRASGGG 14
DB 220 GGGSGGSGGGGGG 233

RESULT 11
ID 090KM9 PRELIMINARY; PRT; 306 AA.
AC 090KM9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN, ALTERNATE TRANSCRIPT.
GN RALY.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=99431566; PubMed=10500250;
RA Khrebukova I., Kuklin A., Woychik R.P., Michaud E.J.;
RT Biochim. Biophys. Acta 1447:107-112(1999).
RL EMBL; AF148457; AAF04487.1; -.
DR INTERPRO: IPR000504; -.
DR PFAM: PF00076; trm; 1.
DR PROSITE: PS00030; RNP_1; UNKNOWN_1.
KW Nucleocapsid; Ribonucleoprotein.
SQ SEQUENCE 306 AA; 32463 MW; 7FA376D3BD8E4728 CRC64;

Query Match 81.5%; Score 66; DB 4; Length 306;
Best Local Similarity 85.7%; Pred. No. 0.46;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGSGGSRASGGG 14
DB 235 GGGSGGSGGGGGG 248

RESULT 12
ID 096754 PRELIMINARY; PRT; 509 AA.
AC 096754;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE INTERMEDIATE FILAMENT PROTEIN E2.

OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7740;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=WHOLE ANIMAL;
RX MEDLINE=99019308; PubMed=9804163;
RA Karabinos A., Riemer D., Erber A., Weber K.;
RT "Homologues of vertebrate type I, II and III intermediate filament
RT (IF) proteins in an invertebrate: the IF multigene family of the
RT cephalochordate Branchiostoma.";
RL FEBS Lett. 437:15-18(1998).
CC -1- SIMILARITY: TO ALL OTHER INTERMEDIATE FILAMENT PROTEINS.

DR EMBL; AJ010293; CA09067.1; -.
DR INTERPRO: IPR001664; -.
DR INTERPRO: IPR002957; -.
DR PFAM: PF00038; filament; 1.
DR PRINTS: PRO1248; TYPEKERATIN.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Heptad repeat pattern.
SQ SEQUENCE 509 AA; 52437 MW; EAF9975729601C6C CRC64;

Query Match 81.5%; Score 66; DB 5; Length 509;
Best Local Similarity 85.7%; Pred. No. 0.76;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGGSGGSRASGGG 14
DB 461 GGGSGGSGGGGGG 474

RESULT 13

045030 ID 045030 PRELIMINARY; PRT; 584 AA.
 AC 045030;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
 DE STRABISNUS.
 GN STBM.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CANTON S.; TISSUE=EYE-ANTENNAL IMAGINAL DISC;
 RL Wolff T., Rubin G.M.;
 RL Development 0:0-0(1998).
 DR EMBL: AF044208; AAC02533.1;
 DR FLYBASE: FBgn0015838; Vang.
 SQ SEQUENCE 584 AA; 66118 MW; 71D94D545C3F1DFA CRC64;

Query Match 81.5%; Score 66; DB 5; Length 584;
 Best Local Similarity 80.0%; Pred. No. 0.87;
 Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGGSGGRASGGGGS 15
 |||||
 DB 26 GGGGSGGGSVSGGGGS 40

RESULT 14
 O9P60 PRELIMINARY; PRT; 592 AA.
 AC O9P60;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE ENDO-1,4-BETA-GLUCANASE.
 GN XF0818.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RC MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvares R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
 RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Gantier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.V., Martins A.M.B.N., Madeira H.M.F., Martino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsushima A.Y.,
 RA Menck C.F.M., Miranda E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nham A.J.T., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Resegues J.B.,
 RA Queirolo R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truifi D., Tsai S.M., Tsunako M.H.,

RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa."
 RL Nature 406:151-157(2000).
 DR EMBL: AE003921; AAF83628.1;
 DR INTERPRO: IPR001547;
 DR INTERPRO: IPR001919;
 DR INTERPRO: IPR002952;
 DR PPM: PR00150; cellulase; 1.
 DR PPM: PR00553; CBD; 2; 1.
 DR PRINTS: PR01228; EGGSHIELD.
 DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
 SQ SEQUENCE 592 AA; 59967 MW; 9846DAEA3B5C89E CRC64;

Query Match 81.5%; Score 66; DB 2; Length 592;
 Best Local Similarity 85.7%; Pred. No. 0.88;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGGGSGGRASGGGGS 14
 |||||
 DB 467 GGGGSGGGSVSGGGG 480

RESULT 15
 O9VK59 PRELIMINARY; PRT; 944 AA.
 AC O9VK59;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE CG5787 PROTEIN (FRAGMENT).
 GN CG5787.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-O., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson R.K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Finkler A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glocker A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Jasti D., Houston K.A., Howland T.J., Wei M.-H., Ilegam C.,
 RA Jostali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshnell A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen F.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Syrakas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003637; AAF53220.1; -
DR FLYBASE: FBgn0032454; CG5787.
FT NON_TER 944 944
SQ SEQUENCE 944 AA; 99435 MM; 96437EBF17FA339F CRC64;

Query Match 81.5%; Score 66; DB 5; Length 944;
Best Local Similarity 78.6%; Pred. No. 1.4;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCGGGGGRASGGG 14
Db 413 GCGGGGGRASGGG 426

Search completed: March 15, 2001, 11:10:56
Job time: 1112 sec